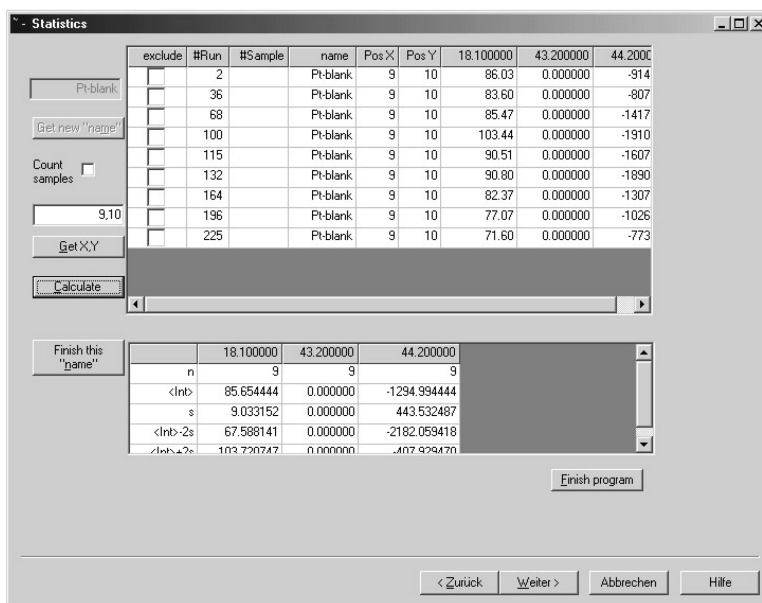


MS-Express: Data-Extracting and -Processing Software for High-Throughput Experimentation with Mass Spectrometry

Pierre-Alain W. Weiss, Christof Thome, and Wilhelm F. Maier

J. Comb. Chem., **2004**, 6 (4), 520-529 • DOI: 10.1021/cc034050m • Publication Date (Web): 30 April 2004

Downloaded from <http://pubs.acs.org> on March 20, 2009



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

[View the Full Text HTML](#)



ACS Publications
 High quality. High impact.

MS-Express: Data-Extracting and -Processing Software for High-Throughput Experimentation with Mass Spectrometry

Pierre-Alain W. Weiss, Christof Thome, and Wilhelm F. Maier*

*Lehrstuhl für Technische Chemie, Universität des Saarlandes, Im Stadtwald,
D-66123 Saarbrücken, Germany*

Received October 1, 2003

High-throughput experiments (HTE) result in large amounts of raw data that have to be evaluated for sample classification. Especially mass spectrometry, a widely used detection method in catalytic HTE applications, produces enormous amounts of data. In the past few years, in catalysts research, several test rigs based on mass spectrometric detection have been independently reported by different groups. In a typical HTE, the catalysts are tested sequentially; the recording of the scans, however, occurs continuously. For this reason, the scans of interest have to be extracted from the raw data, and scans belonging to the same sample have to be averaged in a tedious procedure before further processing. In this publication, we present our custom-designed software MS-Express (*mass spectrometry data-extracting and -processing software*), an efficient tool for HTE MS data evaluation. MS-Express not only sorts the data, it also establishes statistical significance with the help of reference and blank data and provides concise information about abundance and intensity distributions of expected peaks. A special feature is that the program also reports unexpected MS signals, which potentially lead to unexpected discoveries.

Introduction

Time and cost-saving high-throughput techniques are increasingly applied for the discovery and optimization of materials and catalysts.¹ Efficient high-throughput experimentation requires a complete combinatorial workflow, such as library design tools, automated synthesis, a technology for the high-throughput screening (HTE) of desired properties, custom-designed software for the evaluation and visualization of HTE data, and data transfer and storage.² Such complete setups are presently only available in a few industrial laboratories, and the development of individual building blocks and their application is typical for the present research activities.

In the recent development of new high-throughput methods for the screening of solid catalysis, special emphasis has been placed on mass spectrometry (MS) detection. Chen has successfully applied modified electrospray ionization tandem mass spectrometry (ESI-MS) for the screening of homogeneous catalysts in polymerization reactions, such as olefin metathesis or ethylene polymerization.³ Symyx has developed custom-designed MS equipment to monitor product composition of heterogeneously catalyzed reactions sequentially by QMS (quadrupole mass spectrometry) in primary screening applications.⁴ The catalysts adhering to a flat library are heated from the back by a CO₂ laser in a pressure chamber while the catalyst sample is supplied with the feed gas mixture through a capillary on the opposite side. Simultaneously, a small part of the feed-product mixture is transported by another capillary to the MS for detection. Senkan et al. presented a simpler secondary screening device.⁵ Catalysts prepared by impregnation of prefabricated silica or alumina tablets are manually placed into individual

channels of a heated gas-phase flow reactor with multiple channels in a ceramic block. The whole reactor is moved by an xyz table to insert the fixed capillary of the MS into the exit opening of individual reactor outlets for sequential sampling. In another typical secondary screening setup, a multitubular reactor with 15 parallel gas-phase flow quartz tubes has been reported by Claus et al.⁶ The product gas is sequentially switched to a QMS by a 16-valve multiport for sampling. In a miniaturized approach, the same group reported the use of the multiple channel arrays of a micro-mixer in combination with QMS for the HTE of solid catalysts.⁷ While in all these approaches closed reactors have been applied, our primary screening setup has been based on open library plates consisting of several hundred wells in which individual catalysts have been placed.⁸ A capillary bundle comprising a feed gas capillary and a capillary for product gas transfer to the QMS is moved into the open wells sequentially. Two setups have been realized: one moves the capillary bundle with a modified pipetting robot,^{8,9} the other moves the reactor with the help of an xyz table.¹⁰ Such an open reactor system is a compromise to obtain high-throughput. It is convenient to use and simple, but O₂ diffusion to the catalyst surface is unavoidable, and the data are not as reliable as those obtained in a closed reactor system. It was surprising to note that even hydrogenation reactions can be carried out successfully in such open reactors.

All of these different high-throughput reactors based on MS analysis have in common a rapid accumulation of data, which requires the efficient help of automated data evaluation. Manual inspection of the data accumulated renders the experimental high-throughput approach obsolete. Until now, none of the groups involved have presented their solution

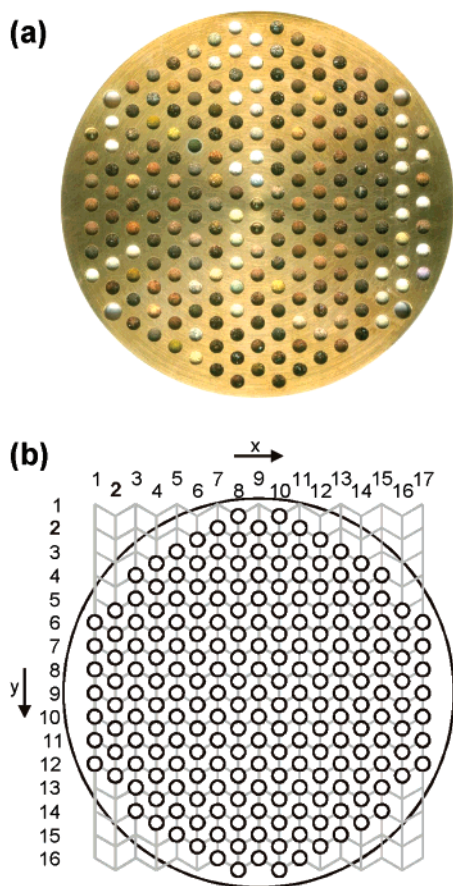


Figure 1. Library plate with 207 positions: (a) photography; (b) x/y position labeling.

for the automated data handling and evaluation resulting from a HTE with MS.

Results and Discussion

A typical MS-based HTE starts with the selection of the testing conditions and the design of the catalyst library. The catalyst preparation procedures selected are sol-gel-based, which relies on the mixing of liquid reagents and solutions and allows the use of dispensing robots.

The design of the libraries is realized in tailored software called Plattenbau.¹⁰ While the user provides information on library layout and design, doping, or composition spreads, Plattenbau calculates the volumes to dispense from available recipes and creates and exports optimized dispensing lists that can be used by the software of the dispensing robots. Samples are synthesized directly either in the wells of a library plate or into arrays of flasks by reagent dispensing with the dispensing robot. If synthesized directly into the library plate, the resulting gels are dried and calcined in the plate before use. If arrays of flasks are used for synthesis, the materials are dried and calcined, and the resulting powders are then milled and filled into the plate wells manually, which has the advantage of comparable fill heights and texture of the catalyst bed. Figure 1 shows a typical library plate ($\varnothing = 99$ mm) with 207 wells and the x/y labeling used for identification of the well and catalysts positions. This particular library plate will also serve for further explanations in this publication. It comprises 204 powder

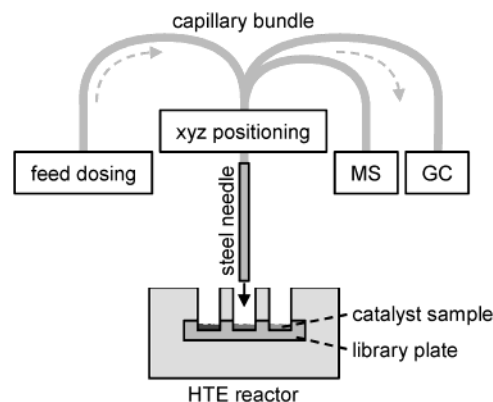


Figure 2. Scheme of our setup for high-throughput screening of solid catalysts.

samples consisting of platinum-doped oxides and mixed oxides, an empty well (blank) in position 9/9, a reference catalyst (reference) in 1/6, and a well containing only the platinum doping and no oxide (Pt blank) in 9/10.

The catalytic behavior of the library plate described above has been tested for the hydroisomerization of *n*-pentane at 300 °C in our high-throughput test rig with MS and gas chromatography (GC) detection.⁹ The setup of the HTE is schematically shown in Figure 2. The catalysts are present as solids on the bottom of the wells of the library plates. The central part of the setup is a capillary bundle ending in a steel needle, which is automatically positioned by a modified synthesis robot into the library wells for reaction product sampling. All capillaries are heated by heating wires and are thermally insulated to avoid condensation of feed and products. The gaseous and liquid reagents, a gaseous mixture of *n*-pentane and hydrogen in this case, enter the feed gas capillary via a preheating and mixing chamber. A ceramic mask covers the library plate, thus transforming the library into an array of open micro reaction chambers with the catalysts sitting at the bottom of the wells. The preheated feed gas overflows the catalyst sample in a reaction chamber and displaces the air in the chamber. The resulting gas mixture containing educts and products are sampled continuously for analysis by additional capillaries connected to a double-focusing sector-field MS (AMD-MMH1) and a GC (Agilent 6890) with a gas sample loop. Additional oxidation reactions competing with the hydroisomerization of *n*-pentane are probable because a small amount of oxygen is always present in the reaction chamber through diffusion. For the experiment discussed here, the isomers *n*-pentane and *i*-pentane were detected with the GC, and the MS was used for the detection of oxidation products such as CO₂ and H₂O. The overall measurement time used was 80 s per well.

For the control of a HTE experiment with the setup described above, a custom-designed software called TestRig has been developed.^{9,10} TestRig controls the movement of the steel needle, permitting screening of the library plate in an arbitrary sequence, and sends a start signal to the GC after entry of the needle in a reaction chamber. Individual library positions may be measured several times, for example, to repeat blank measurements between the sample measurements. The sequence to be measured is easily edited and is saved in ASCII format in a file describing the desired

(a)		(b)	
Art	:Hexagonal	2002-08-20	15:55:18.170 ; in ; 9 ; 9
Rackgroesse X=	17	2002-08-20	15:56:20.620 ; out ; 9 ; 9
Rackgroesse Y=	16	2002-08-20	15:56:40.340 ; in ; 9 ; 10
1.)	X= 9 Y= 9	2002-08-20	15:57:42.840 ; out ; 9 ; 10
2.)	X= 9 Y= 10	2002-08-20	15:58:04.650 ; in ; 1 ; 6
3.)	X= 1 Y= 6	2002-08-20	15:59:07.100 ; out ; 1 ; 6
4.)	X= 9 Y= 9	2002-08-20	15:59:28.690 ; in ; 9 ; 9
5.)	X= 9 Y= 9	2002-08-20	16:00:30.420 ; out ; 9 ; 9
6.)	X= 1 Y= 7	2002-08-20	16:00:49.920 ; in ; 9 ; 9
7.)	X= 1 Y= 8	2002-08-20	16:01:51.600 ; out ; 9 ; 9
8.)	X= 1 Y= 9	2002-08-20	16:02:12.800 ; in ; 1 ; 7
9.)	X= 1 Y= 10	2002-08-20	16:03:14.870 ; out ; 1 ; 7
10.)	X= 1 Y= 11	2002-08-20	16:03:34.640 ; in ; 1 ; 8
		2002-08-20	16:04:37.260 ; out ; 1 ; 8
		2002-08-20	16:04:56.980 ; in ; 1 ; 9
		2002-08-20	16:05:59.540 ; out ; 1 ; 9
		2002-08-20	16:06:19.310 ; in ; 1 ; 10
		2002-08-20	16:07:21.980 ; out ; 1 ; 10
		2002-08-20	16:07:41.750 ; in ; 1 ; 11
		2002-08-20	16:08:44.480 ; out ; 1 ; 11

Figure 3. Cutout of the (a) desired and (b) measured sequence lists exported from TestRig.

sequence. The status of the effectively measured sequence is permanently saved in another ASCII file containing entering times (t_{in}) and exiting times together with the x and y position coordinates for each measurement. Cutouts of both sequence lists generated by TestRig are shown in Figure 3. The serial number describing the HTE sequence order in the desired sequence list will be referred to as run# in the following. Each run number can easily be assigned to the position of a measured well with the help of both sequence lists.

The software MASPEC Data System (Mass Spectrometry Services) controls the sector-field MS and the MS data registration. The gas sampling to the MS and the recording of the scans occur continuously, and several scans are registered during the testing of one catalyst as well as during the positioning of the capillary bundle and the movements outside of the reactor. The MS raw data of a whole HTE sequence are saved in a single file that contains more than 6000 mass spectra (scans) in a typical HTE. For the evaluation of a HTE with MS, only the scans recorded in the reaction chambers are of interest. These scans have to be extracted from the data file generated by MASPEC and associated with the respective run number and, accordingly, to a well position before further processing (such as averaging the scans for each run number, standardization, blank correction.). This task was accomplished in a tedious manual procedure that took several days in our early stage of development, proving in a painful manner the obvious necessity for an automated data extraction and processing software.

Figure 4 shows the dependency of some ion currents on the scan number (scan#). This presentation clarifies the changes in signal intensity from one scan to the next and thus to follow the actual progress of the HTE. The intensity at $m/z = 43$ predominantly originates from the educt n -pentane. The intensity at $m/z = 44$ is composed of the oxidation product CO_2 and of hydrocarbon fragments. By watching the profile, it is easy to recognize whether a scan was recorded in a reaction chamber or in air during the steel needle movement outside of the reactor. After entering a reaction chamber, the feed gas partly replaces the present

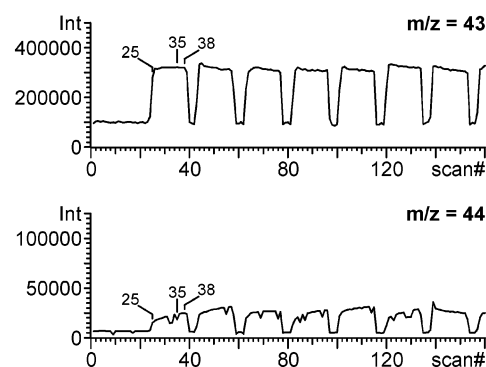


Figure 4. Selected ion traces from a typical screening with MS (single scans are highlighted).

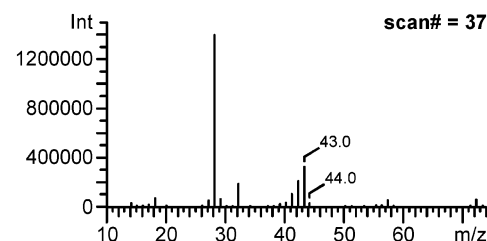


Figure 5. Representative scan measured in a blank position.

air, leading to an increase of the signal $m/z = 43$. It is obvious from this presentation that the scans of interest have to be extracted from the raw data for evaluation, and scans belonging to the same sample have to be averaged in a tedious procedure before further processing.

The MS monitoring of catalytic reactions with hydrocarbons leads to a large number of ion fragments and, consequently, to considerable MS data. In some cases, more than 200 signals were recorded in a single scan, leading to very large exported data sets. Every MS signal is defined by three coordinates: its serial scan number (scan#), its mass to charge ratio (m/z), and its corresponding intensity (Int). A representative scan (scan# = 37) recorded during a blank measurement is pictured in Figure 5. The fragments shown result from the n -pentane/ H_2 feed gas and from the air diffusing into the reaction chamber. A file containing all MS raw data of an HTE can be exported by MASPEC in ASCII format. A segment of this file is shown in Figure 6 for

```

;Scan Number: 37, Time : 0:02:40.77.
;Number of Peaks=33.
;Base M/z=28.2, 100% Int.=1390274.
;
; ---M/z--- ---Int.--- --Width- ---%age---
;
  14.1      25738      16      1.85
  15.1      1839      15      0.13
  16.1      4346      17      0.31
  17.1      11402     16      0.82
  18.2      63045     16      4.53
  20.1      1825      15      0.13
  26.1      2596      16      0.19
  27.2      44015     15      3.17
  28.2     1390274     15     100.00
  29.2      57222     23      4.12
  30.2      1046      15      0.08
  31.2       297       9      0.02
  32.2     183822     15     13.22
  34.2       741      13      0.05
  37.2       586      12      0.04
  38.2      1644      15      0.12
  39.2     18588      15      1.34
  40.2     26458      19      1.90
  40.2     2552      10      0.18
  41.2     99721      16      7.17
  42.3     202753     15     14.58
  43.3     319781     16     23.00
  44.2     24710      27      1.78
  50.2       715      13      0.05
  51.3       992      14      0.07
  53.4     1594      15      0.11
  55.4     7312      15      0.53
  56.4     8732      16      0.63
  57.4     48491     16      3.49
  58.4     2182      15      0.16
  71.3     2489      15      0.18
  72.3     53490     15      3.85
  73.4     3015      15      0.22
#End Scan: 0:02:40.77

```

Figure 6. Cutout of the scan text file exported from MASPEC for scan 37 (m = multiplet).

scan# = 37. The data are primarily sorted by their scan number, leading to a consecutive listing of all scans in tables containing the m/z -Int data pairs (scan# \leftarrow m/z -Int). The user, however, is interested in the devolution of a m/z value of interest with the chronological progress of the HTE, that is, with increasing scan number. Consequently, all data must be sorted by their mass-to-charge ratios ($m/z \leftarrow$ scan#-Int). This sorting can be done with MASPEC, leading to the display of ion currents, that is, the chronological devolutions of the intensities for fixed m/z values. The sorting with MASPEC is only possible for preselected m/z values, which means that one must decide which m/z values are of interest for the evaluation of the catalytic screening before knowing the results. The signals generally chosen under these conditions are those of few educt and main product fragments. In such an evaluation strategy, interesting catalysts with high activity for the formation of a product different from the expected one will not be recognized. How this can be overcome is discussed below.

In the following, our custom-designed software, MS-Express, which treats all the problems just discussed, is described.

MS-Express. The software MS-Express for high-throughput data MS analysis was developed for the extraction and processing of as many signals as desired. MS-Express (*mass spectrometry data-extracting and -processing software*) was programmed in Visual C++ with Microsoft Foundation Classes. The data treatment is done in consecutive steps that

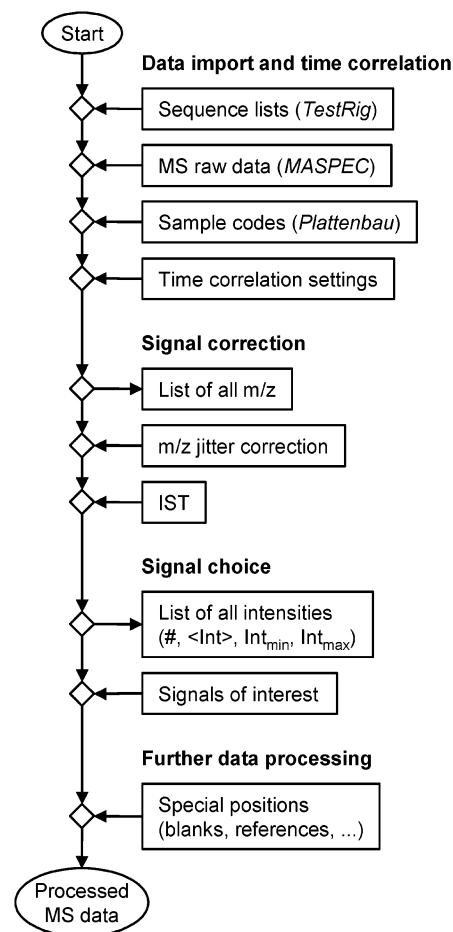


Figure 7. Operating sequence with MS-Express.

are grouped in different windows of a graphical user interface for ease of use. This structuring of MS-Express is depicted in the overview of the operating sequence in Figure 7. The workflow starts with the import of all necessary data from TestRig and MASPEC and basic settings for the correlation of information from different files. A list containing the m/z values of all signals detected during the HTE is displayed, small fluctuations in the m/z value for the same ion fragment from one scan to another or as a result from an instrumental drift are corrected automatically, and an internal standard (IST) used for intensity correction can be set (optional). After standardization, the signals belonging to the same run number are averaged, resulting in an overview that contains summarized information about the data set. This simplifies the selection of signals of interest. The next steps in the software are associated with the involvement of special positions, such as blanks or references into the data treatment and respective correction of intensities. A more detailed description of all functions is given below.

Data Import and Time Correlation. MS-Express requires information from two programs, the timing from the software controlling the sequence of the experiment (here TestRig) and the MS data from the Software controlling the mass spectrometer (here MASPEC). It starts by prompting the user to load up for evaluation several files resulting from the HTE (Figure 8): (1) the desired sequence list from TestRig (see Figure 3a), (2) the measured sequence list from TestRig (see Figure 3b), (3) the MS data file from MASPEC

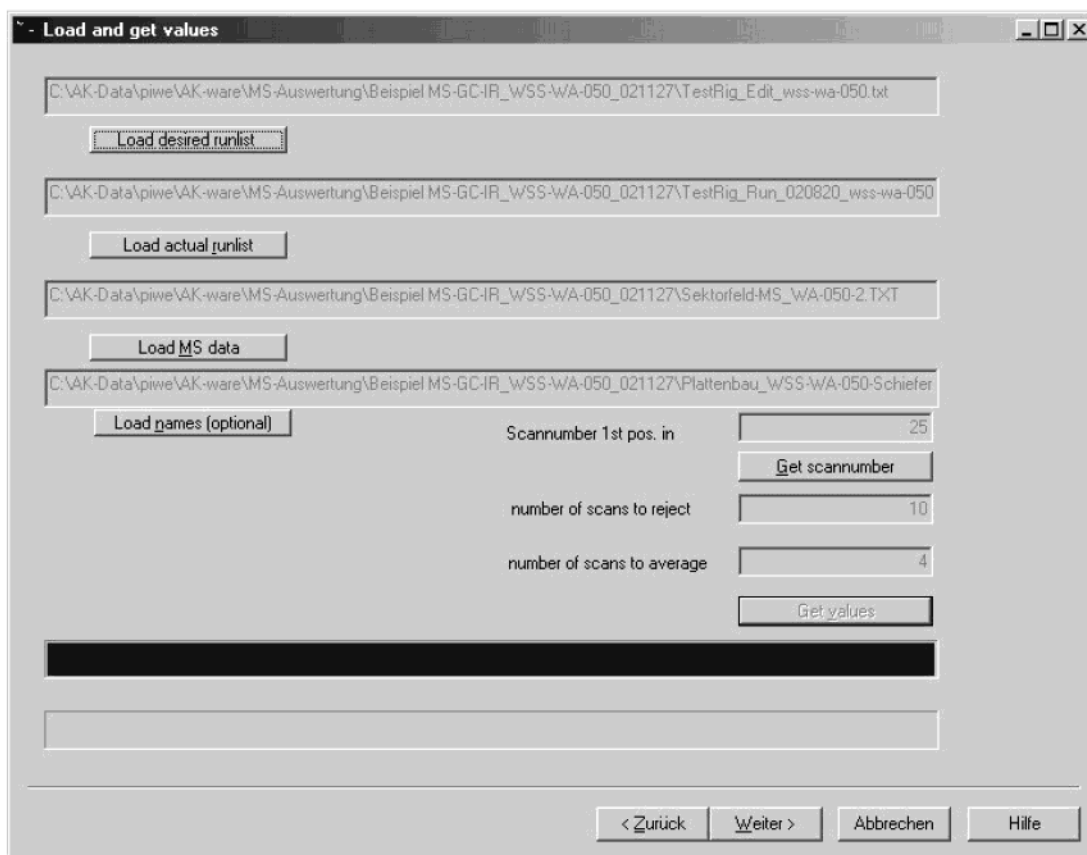


Figure 8. Data import and time correlation in MS-Express.

(see Figure 6) and (4) optionally an XML file generated by Plattenbau that contains a brief name for each sample describing its composition (for example, Si100 for platinum-doped silicon oxide, Si₅₀Ti₅₀ for a platinum-doped mixed oxide). The program continues only if the entries in the desired sequence list correspond to those of the measured sequence list to prevent possible misinterpretations as a consequence of last minute changes in the sequence order. A shorter measured sequence list resulting from a possible preterm abort of a HTE is accepted, however. The resuming of both sequence lists leads to the attribution of the times recorded by TestRig at every entrance into a reaction chamber (t_{in}) to its associated run number.

The MS data file contains a consecutive listing of all scans sorted by their serial scan number in reference to their recording time (t_{scan}). These relative time values are set by an internal clock in MASPEC that starts with $t_{scan} = 0$ at the manual actuation of the MS detection in MASPEC that generally occurs shortly before the sequence start in TestRig. To correlate the time axes in MASPEC and TestRig, it is necessary to identify the first scan monitored after entrance of the steel needle in the first well of the sequence (run# = 1). This value is easily determined with the first drastic change in intensity in the ion currents and is entered in MS-Express that correlates both time axes. The next step is the assignment of all scans to evaluate to their respective run number. This is done by once typing in two additional values: first, the number of scans recorded after entrance of the steel needle that should not be used for evaluation; and second, the following number of scans recorded for the same run# that should be used for evaluation and, conse-

quently, should be averaged. The entrance of the steel needle into a well is recognized by the increase and the exit by the decrease in the intensity of $m/z = 43$ (see Figure 4). The associated scan numbers are used for the correlation of the time axes from the TestRig and the MASPEC files. During the time of the well sampling, several MS scans, depending on MS settings, are recorded (here, 14 scans). It is advisable to reject some of the initial scans and chose the scans toward the end of the sampling for averaging (here, we have chosen to reject the first 10 scans and average the final 4 scans for evaluation). After the assignment of the selected scans to their run number, MS-Express deletes unnecessary additional scans, thus reducing the temporary data set. The averaging of the respective scans is done at a later point in the data processing.

MS-Express is not limited to mass spectrometric files generated by MASPEC. An import tool has been implemented that also permits the use of files generated by QUADSTAR (Inficon), the software controlling a quadrupole mass spectrometer in our second high-throughput test rig with MS detection.¹⁰

Signal Correction. After the designation of the scans to evaluate, all m/z values in the data set are listed on the screen (Figure 9). In the low-resolution detection mode used with the sector-field MS, an identical fragment may lead to slightly different m/z values from one scan to another due to small fluctuations in the initial positioning of the scanning magnet. For example, the signal of the main water fragment appeared at $m/z = 18.1$ as well as at $m/z = 18.2$ in this experiment. To combine these signals, a special m/z jitter correction using the difference between two m/z values can be applied. In a

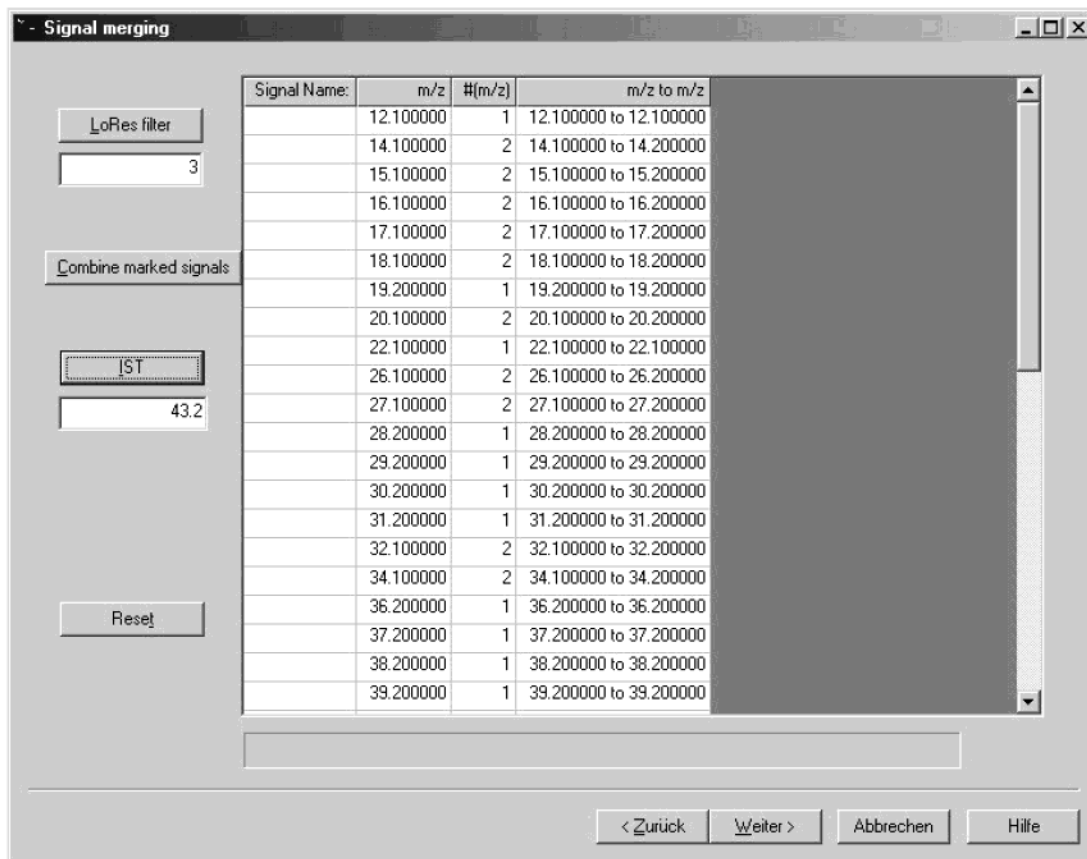


Figure 9. Signal merging and standardization in MS-Express.

low resolution measurement, MASPEC delivers m/z values with an accuracy of one position after the decimal point. Accordingly, the difference of two values belonging together that directly follow each other can only be 0.1. Signals with m/z values lying farther apart do not belong to the same fragment. MS-Express combines all m/z values that have a difference of $\Delta(m/z) = 0.1$ from the preceding value. If a signal appears at several m/z values in the same scan series (false multiplet), the intensities are simply added. The maximum amount of consecutive m/z values to combine is set in the graphical interface as a filter parameter, whose value is a compromise. On one hand, a value that is too large may lead to the combination of different fragments; a smaller value may be insufficient for a reliable automated correction. An adequate value found empirically for our experiments is 3. This filter parameter setting causes the combination of m/z values with a variation of ± 0.1 after pressing the “LoRes filter” button. The column titled “#(m/z)” displays the quantity of m/z values combined in the new artificial signal. The m/z range combined in the new signal is shown on the right column. The user can still manually combine the combination of m/z values belonging together that were not recognized by the m/z jitter filter. For this purpose, the user marks the respective m/z values in the table and unites them by pressing a second button (“combine marked signals”). Even if combining different signals may look unusual to advanced mass spectrometry users, this jitter correction does not create a signal mismatch here. In fact, it is absolutely necessary for an automated evaluation of data sets with slightly fluctuating or drifting m/z values.

In general, fluctuation of the signal intensities due to small temporary changes in the detection sensitivity can be minimized with an internal standard (IST). The fragment chosen for standardization should be representative of the chronological progress of the detection sensitivity over the whole experiment. This is ideally a fragment of a molecule contained in the feed gas that does not participate in the reaction. The user optionally selects the m/z value of an IST in the table and enters it into the designated edit box of the user interface. Hence, a potential combination of fluctuating m/z values with the m/z jitter filter has to occur before the standardization. Pressing the IST button starts the standardization. We usually add argon to the feed gas as an IST in MS screenings. Because Ar gas was missing in the feed gas mixture shown here, the intensities of all m/z values were related to the intensities of the signals at $m/z = 43$, which originate from fragments of both pentane isomers and other hydrocarbons, if present. The resulting normalized ratios are related to the first scan selected for evaluation (here, scan# = 35) by multiplication

$$\frac{\text{Int}_{m/z=i;\text{scan}=j}}{\text{Int}_{m/z=43;\text{scan}=j}} \times \text{Int}_{m/z=43;\text{scan}=35}$$

The assignment of single scans to their respective run number by correlation of the time axes has been discussed in the previous section. Because the averaging function used for the assignment of the mean intensities to their respective run number contains a sum, this assignment can only be done after the optional standardization. It is started automatically

Evaluate	m/z	m/z to m/z	# holes	<Int>	max. Int[%]	min. Int[%]
<input type="checkbox"/>	27.100000	27.100000 to 27.200000	228	45348.334999	105.31	97.84
<input type="checkbox"/>	28.200000	28.200000 to 28.200000	228	1329827.796943	114.90	80.86
<input type="checkbox"/>	29.200000	29.200000 to 29.200000	228	57965.508651	123.62	95.45
<input type="checkbox"/>	30.200000	30.200000 to 30.200000	228	1097.105551	158.88	89.94
<input type="checkbox"/>	31.200000	31.200000 to 31.200000	164	145.948009	292.53	40.64
<input type="checkbox"/>	32.100000	32.100000 to 32.200000	228	97726.917057	234.40	31.84
<input type="checkbox"/>	34.100000	34.100000 to 34.200000	81	570.228982	169.02	10.86
<input type="checkbox"/>	36.200000	36.200000 to 36.200000	2	328.949406	102.48	97.52
<input type="checkbox"/>	37.200000	37.200000 to 37.200000	228	553.014475	163.38	74.72
<input type="checkbox"/>	38.200000	38.200000 to 38.200000	228	1717.903699	135.96	94.33
<input type="checkbox"/>	39.200000	39.200000 to 39.200000	228	19378.144980	111.80	96.44
<input type="checkbox"/>	40.100000	40.100000 to 40.300000	228	27467.814918	114.27	81.62
<input type="checkbox"/>	41.200000	41.200000 to 41.300000	228	101910.638426	104.57	97.79
<input type="checkbox"/>	42.200000	42.200000 to 42.300000	228	208632.665798	103.34	97.24
<input checked="" type="checkbox"/>	43.200000	43.200000 to 43.300000	228	322830.000000	100.00	100.00
<input checked="" type="checkbox"/>	44.200000	44.200000 to 44.300000	228	31154.819696	209.70	69.74
<input type="checkbox"/>	45.200000	45.200000 to 45.200000	5	390.825777	196.00	16.31
<input type="checkbox"/>	50.200000	50.200000 to 50.300000	228	742.455352	138.57	91.57
<input type="checkbox"/>	51.200000	51.200000 to 51.300000	228	1006.643302	134.83	90.97

Figure 10. Choice of signals of interest in MS-Express.

by MS-Express before moving to the next program window, thereby drastically reducing the data set. For more clarity, the calculation for $m/z = 44$ in run# = 1 is given here.

$$\frac{1}{4} [\text{Int}_{m/z=44;\text{scan}=35} + \text{Int}_{m/z=44;\text{scan}=36} + \text{Int}_{m/z=44;\text{scan}=37} + \text{Int}_{m/z=44;\text{scan}=38}]$$

Signal Choice. In the next window, a tabular overview of the complete HTE is displayed (Figure 10). The signals are listed with increasing m/z . The column “# holes” shows how often a signal was found. Because the HTE sequence discussed here consisted of 228 single measurements (run# = 1–228), the readout 228 at $m/z = 44.2$ signifies that the respective fragment was detected in every run number. The column <Int> displays the mean of the intensities over all run numbers in which the signal was detected. The columns “max. Int[%]” and “min. Int [%]” show the relative extremes of the signal’s intensity based on the respective mean intensity. No extreme is reported for $m/z = 43.2$, because this mass signal was used as IST. It is obvious at the first glance that the CO_2 product signal at $m/z = 44.2$ (<Int>) is far above the system noise and that there may be catalytically active wells producing much more CO_2 than the mean of all samples (max. Int [%]). The user may now decide whether this signal is of interest to him or if other fragments of the same product are even more appropriate for evaluation. The signal is accepted for evaluation by checking this line in the left column. Thus, evaluation is no longer limited to a few preselected educt and product signals, but all results of the HTE are rapidly judged. This overview is one of the most

valuable features of MS-Express because it permits the fast recognition but unexpected mass signals that could originate from a scientifically or economically significant or unexpected catalytic product (enhancement of discovery chances).

Further Data Processing. The next steps in the workflow are directed to library positions that contain various reference samples and are measured repeatedly during the HTE. Positions belonging to this class are blank, reference, and further special positions useful for comparison. The user is guided through a program window for the processing of blank positions first, followed by an analogous window for reference samples and another one for additional special positions. Because the windows are very similar, the functions are described with the help of the screenshot in Figure 11. The user enters a name for the special position to process in the upper left edit box (“Pt blank” in this case) and confirms it by actuating the button below (“Get new ‘name’”). In the preceding steps, the name “Blank” is automatically attributed to blank positions; the name “Reference,” to reference positions. The x/y position coordinates of all wells filled with “Pt blank” are successively entered into the edit box above the button “Get X, Y” and confirmed by its activation. Here, the only Pt blank has been placed in position 9/10. Clicking “Calculate” informs the software that no further positions will follow and results in the display of all signals of interest detected in the indicated positions. The user can exclude a measurement from the subsequent calculation of the average and of the standard deviation by checking the respective box. The results are displayed in the lower table of the window. Pressing “Finish this ‘name’”

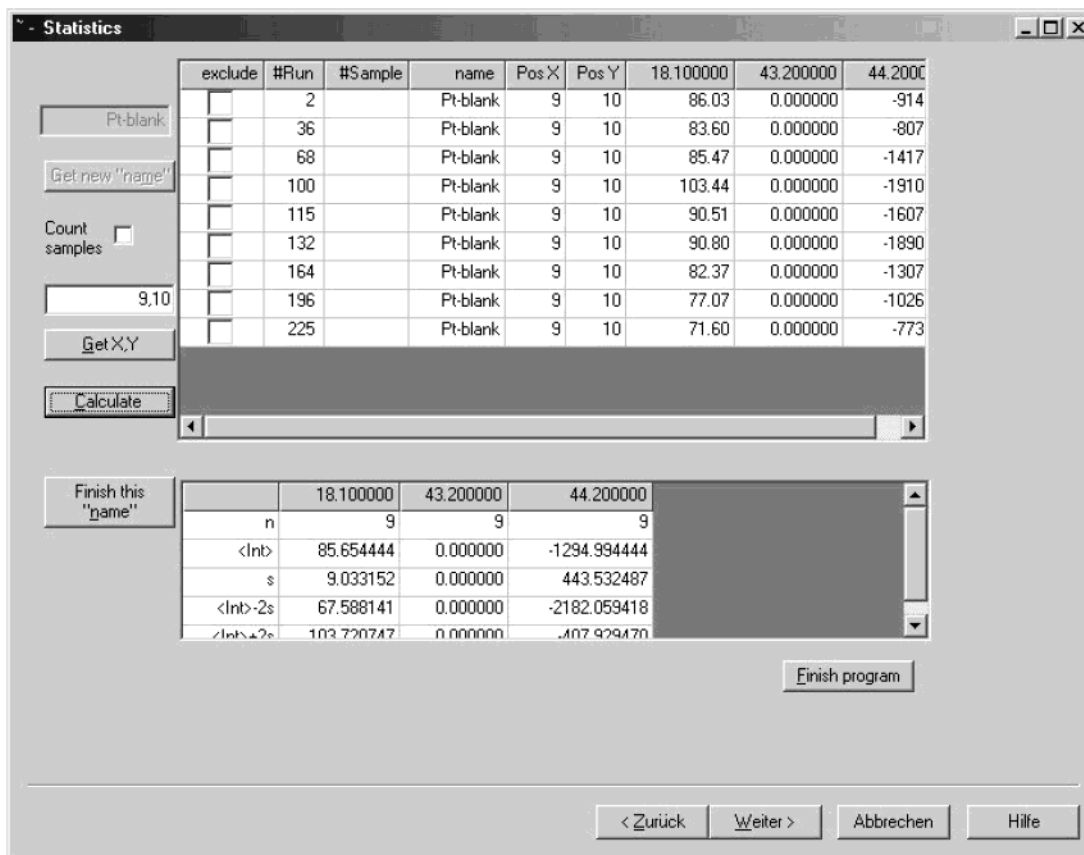


Figure 11. Further processing in MS-Express: blank corrections, reference normalization.

ends the input for “Pt blank”, and the same process can be repeated for other samples.

MS-Express proceeds in an analogous way for blank and reference positions. If the library plate screened does not contain blank or reference positions, the associated steps can be skipped. With the intensities obtained in the selected blank measurements ($\text{Int}_{m/z=i;\text{run}\#}$) the mean value ($\langle \text{Int} \rangle_{m/z=i;\text{blank}}$) and the standard deviation ($s(\text{Int})_{m/z=i;\text{blank}}$) are determined. For background correction, the mean value is subtracted from all signals.

$$\text{Int}_{m/z=i;\text{run}\#} - \langle \text{Int} \rangle_{m/z=i;\text{blank}}$$

Only samples with a background corrected intensity larger than

$$\langle \text{Int} \rangle_{m/z=i;\text{blank}} + 2[s(\text{Int})_{m/z=i;\text{blank}}]$$

are be regarded as statistically significantly active. The probability that a value fulfilling this criterion is a true catalytic activity and not an artifact by background noise equals 97.7% if the respective sample is measured only once. Following normalizing with the mean of the reference measurements ($\langle \text{Int} \rangle_{m/z=i;\text{reference}}$) leads to the relative activities.

$$\frac{\text{Int}_{m/z=i;\text{run}\#}}{\text{Int}_{m/z=i;\text{reference}}} \times 100\%$$

This calculation makes sense only if the catalytic activity and selectivity of the reference samples are highly reproducible from one HTE to the next.

The goal of an HTE and of the data processing described here is to assign a value representative for a product yield to the catalyst samples. Blank, reference, and other special measurements of lower interest are solely necessary for a more reliable data processing. They are dispensable in the graphical presentation of the catalysts' results and should be easily extracted from the data set for this reason. Hence, in addition to the run number, a position number (pos#) is introduced that is assigned only once to each position measured in a HTE sequence. At first, just library positions containing real catalyst samples are numbered with a consecutive value in a vertical order, starting at position 1/6 and ending at 17/12. All other positions are added afterward. By sorting the data by their position number, those positions move to the end of the data list and can easily be distinguished from the real samples and separated if desired. The position number is assigned independently of the library layout and thus differs from the x/y position labels. The user decides whether the special positions entered in Figure 11 should be regarded as real catalyst samples by checking the appropriate box after typing in the respective name. The position numbers are not assigned until the end of the automated processing with MS-Express, when it is clear that no additional special position containing a real sample will follow. Actuating the “Finish program” button terminates the software. By definition, the positions containing real catalyst samples are numbered with their position numbers as described first; followed by the blank positions; then by the reference positions; and finally, by all the remaining special positions. Accordingly, here, the catalyst sample in

Run#	Pos#	x	y	18.1	43.2	44.2	Name	XML-Name
6	1	1	7	51322	0	1071		Si100
7	2	1	8	105589	0	-984		Ti100
8	3	1	9	90185	0	1843		Sn100
9	4	1	10	83985	0	-1901		Ta100
10	5	1	11	-14382	0	380		Te100
11	6	1	12	51322	0	1071		W100
12	7	2	5	105589	0	-984		Al100
13	8	2	6	90185	0	1843		Zr100
14	9	2	7	83985	0	-1901		Si50Ti50
15	10	2	8	-14382	0	380		Si50Sn50

Figure 12. Part of a data sheet with blank corrected intensities generated by MS-Express (Data_bla.txt).

Table 1. Files Generated by MS-Express

file name	content
Overview.txt	protocol with description of the six files specified in this table settings used for time correlation overview of all m/z signals after m/z jitter correction; signals selected for evaluation are labeled
Data_raw.txt	processed data with raw intensities
Data_ist.txt	processed data after standardization with IST
Data_bla.txt	processed data after blank correction
Data_nor.txt	processed data after normalizing with references
Statisti.txt	results of statistical calculations with blank, reference, and other special positions

position 1/7 gets the pos# = 1, the last catalyst sample in 17/12 the pos# = 204, the blank in 9/9 the pos# = 205, the reference catalyst in (1/6) the Pos# = 206, and the special position "Pt blank" in 9/10 the pos# = 207.

Output and Presentation of Data Processed. After the data processing, MS-Express creates a folder "MS_Data" containing six files that are described in Table 1. The processed data are exported in an ASCII format that is easily imported into databases. A cutout of the file Data_bla.txt containing the data sheet with IST standardized and blank corrected intensities is shown in Figure 12. From left to right, the columns contain the following information: the run number describing the order of the screening sequence; the position number used for easy selection of true catalyst sample measurements; the x and y position coordinates; the intensities for the m/z values indicated in the title; the names generated in MS-Express for special positions (for example, Pt blank); and finally, the sample names generated in Plattenbau for the synthesis of this library.

The automated presentation of the data listed in the resulting ASCII files is achieved easily in commercial software such as Excel (Microsoft), Origin (Microcal), or Spotfire (Spotfire) with the help of templates and macros. The dependence of the intensities in the blank measurements on the chronological run number is displayed in Figure 13a for $m/z = 44.2$. Two-dimensional presentations of the data in their chronological order are useful for the recognition of possible trends over the HTE sequence, because they may result from a cooling of the reactor or a feed gas loss, for example. The intensities of the catalyst samples are shown in Figure 13b. Only intensities with values larger than 7045 should be considered as active. The intensities of $m/z = 44.2$ are shown independent of their position on the library plate

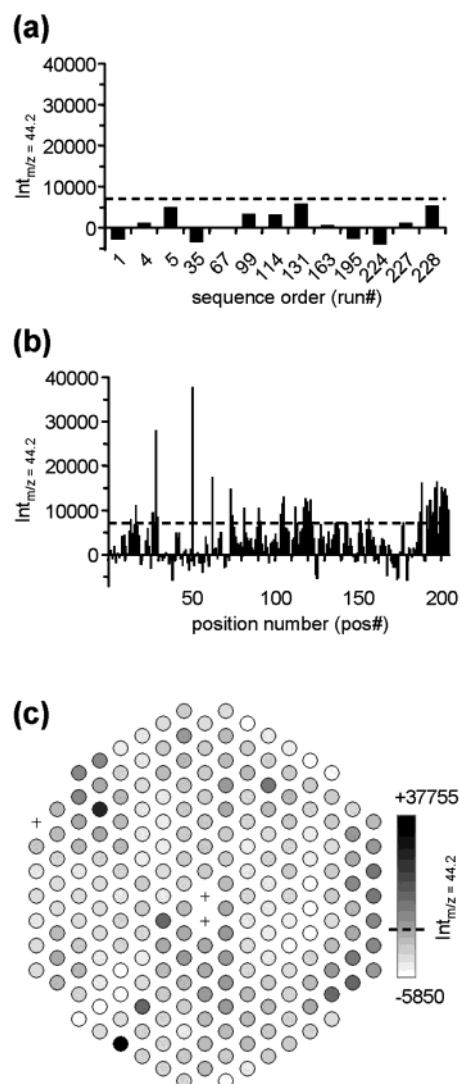


Figure 13. (a) Two-dimensional presentation of the blank measurements (in pos# = 205) in chronological order and (b) of the screening results for catalyst samples; (c) three-dimensional visualization of the screening results for catalysts samples. The dashed lines indicates the limit for statistically significant intensities: $\langle \text{Int} \rangle_{m/z=i;\text{blank}} + 2[s(\text{Int})_{m/z=i;\text{blank}}] = 7045$.

in Figure 13c. The intensity of a sample is specified by the color. The dark spots here identify the two most active catalysts on this library for the combustion of n -pentane. Possible trends over the library plate that could result from a temperature gradient, for example, are readily identified in such a presentation. With the help of this visualization strategy, active catalysts are rapidly identified on the screen. The complete data processing and visualization procedure

described in this example takes only ~15 min for all data of a HTE on a PC running NT4.0 with a 1-GHz processor.

Summary

The automated evaluation of a typical catalytic screening experiment with high-throughput mass spectrometry has been presented. The program module MS-Express can be adapted for the evaluation of data produced by further mass spectrometric controlling software and is expected to drastically accelerate high-throughput research in heterogeneous catalysis and other research fields. The time necessary for a complete MS data evaluation and visualization gained from a sequence with more than 200 catalysts is reduced from several days to approximately 15 min with the procedure described. The time saved is not the only advantage of the automated evaluation. The software presents all data in a summarizing overview for facile identification of signals of interest. A special strength of this software package is that it does not miss interesting unexpected byproducts contained in the data set. Such unexpected discoveries are easily detected by the user in the final output.

The automated data processing strategy presented here is not restricted to mass spectrometric application. By minor adaptation of the software, data originating from other analytical methods may be used, as well.

For a copy of MS-Express, please contact the corresponding author.

Acknowledgment. We thank the Bundesministerium für Bildung und Forschung (BMBF) for support of this work

through Grant 03C0307. C.T. thanks BMBF Grant 03C0305C for support. We thank J. Scheidtmann for valuable discussions and D. Klär for his assistance with the development of the import tool for QUADSTAR files and for intensive testing.

References and Notes

- (1) Jandeleit, B.; Schäfer, D. J.; Powers, T. S.; Turner, H. W.; Weinberg, W. H. *Angew. Chem., Int. Ed. Engl.* **1999**, *38*, 2494.
- (2) Scheidtmann, J.; Weiß, P. A.; Maier, W. F. *Appl. Catal.*, **2001**, *222*, 79.
- (3) Chen, P. *Angew. Chem., Int. Ed.* **2003**, *42*, 2832–2847.
- (4) Cong, P.; Doolen, R. D.; Fan, Q.; Giaquinta, D. M.; Guan, S.; McFarland, E. W.; Poojary, D. M.; Self, K.; Turner, H. W.; Weinberg, W. H. *Angew. Chem., Int. Ed. Engl.* **1999**, *38*, 483.
- (5) Senkan, S.; Krantz, K.; Ozturk, S.; Zengin, V.; Onal, I. *Angew. Chem., Int. Ed. Engl.* **1999**, *38*, 2794.
- (6) Rodemerck, U.; Ignaszewski, P.; Lucas, M.; Claus, P. *Chem. Eng. Technol.* **2000**, *23*, 413.
- (7) Claus, P.; Hönicke, D.; Zech, T. *Catal. Today* **2001**, *67*, 319.
- (8) Orschel, M.; Klein, J.; Schmidt, H.-W.; Maier, W. F. *Angew. Chem., Int. Ed.* **1999**, *38*, 2791.
- (9) Weiss, P.-A. W.; Saalfrank, J. W.; Scheidtmann, J.; Schmidt, H.-W.; Maier, W. F. In *High Throughput Analysis: A Tool for Combinatorial Materials Science*; Amis, E., Potyrailo, R. A., Eds.; Kluwer Academic Publishers: New York, 2003; Chapter 8, pp 125–153.
- (10) Urschey, J.; Weiss, P.-A. W.; Scheidtmann, J.; Richter, R.; Maier, W. F., *Solid State Sci.* **2003**, *5*, 909.
- (11) Scheidtmann, J.; Saalfrank, J. W.; Maier, W. F. *Stud. Surf. Sci. Catal.* **2003**, *145*, 13.

CC034050M